

DEGREE ASYMPTOTICS WITH RATES FOR PREFERENTIAL ATTACHMENT RANDOM GRAPHS

EROL PEKÖZ, ADRIAN RÖLLIN AND NATHAN ROSS

*Boston University, National University of Singapore and
University of California, Berkeley*

Abstract

We provide rates of convergence to the asymptotic distribution of the (properly scaled) degree of a fixed vertex in two preferential attachment random graph models. Our approach is to show that these distributions are unique fixed points of a certain distributional transformation that allows us to obtain rates of convergence using a new variation of Stein's method. Despite the large literature on these models, there is surprisingly little known about the limiting distributions so we also provide some properties and new representations including an explicit expression for the densities in terms of the confluent hypergeometric function of the second kind.

1 INTRODUCTION

Preferential attachment random graphs are random graphs that evolve by sequentially adding vertices and edges in a random way so that connections to vertices with high degree are favored. Following the publication of Barabási and Albert (1999), there has been an explosion of research surrounding these types of models; the seminal reference in the mathematics literature is Bollobás, Riordan, Spencer, and Tusnády (2001). One of the main results of Bollobás et al. (2001) is a rigorous proof that the degree of a randomly chosen vertex in a particular family of preferential attachment random graph models converges to the Yule-Simon distribution. Corresponding approximation results in total variation for this and related preferential attachment models can be found in Peköz, Röllin, and Ross (2010) and Ford (2009).

Here we study the distribution of the degree of a fixed vertex in two preferential attachment models. In Model 1 we start with a graph G_2 with two vertices labeled one and two with an edge directed from vertex two to vertex one. Given graph G_n , graph G_{n+1} is obtained by adding a vertex

labeled $n + 1$ and adding a single directed edge from this new vertex to a vertex labeled from the set $\{1, \dots, n\}$, where the chance that $n + 1$ connects to vertex i is proportional to the degree of vertex i in G_n . Model 2 is one studied in Bollobás et al. (2001) and allows for self-connecting edges. There, we start with a graph G_1 with a single vertex labeled one and with an edge directed from vertex one to itself. Given graph G_n , graph G_{n+1} is obtained by adding a vertex labeled $n + 1$ and adding a single directed edge from this new vertex to a vertex labeled from the set $\{1, \dots, n + 1\}$, where the chance that $n + 1$ connects to vertex $i \in \{1, \dots, n\}$ is proportional to the degree of vertex i in G_n and the chance that vertex $n + 1$ connects to itself is $1/(2n + 1)$.

Let $W_{n,i}$ be the degree of vertex i in G_n under either of the models above. Our main result is a rate of convergence in the Kolmogorov metric (defined below) of $W_{n,i}/(\mathbb{E}W_{n,i}^2)^{1/2}$ to its distributional limit as $n \rightarrow \infty$. Though the literature on these models is large, there is surprisingly little known about these distributions. The fact that these limits exist for the first model has been shown in Móri (2005) and Backhausz (2011) and the same result for both models can be read from Janson (2006) by relation to a generalized Pólya urn, although the existing descriptions of the limits are not very explicit. The primary tool we use here to characterize the limits and obtain rates of convergence is a new distributional transformation for which the limit distributions are the unique fixed points. This transformation allows us to develop a new variation of Stein's method; we refer to Ross and Peköz (2007) and Chen, Goldstein, and Shao (2011) for introductions to Stein's method.

To formulate our main result, we first define the family of densities

$$\kappa_s(x) = \Gamma(s) \sqrt{\frac{2}{s\pi}} \exp\left(\frac{-x^2}{2s}\right) U\left(s - 1, \frac{1}{2}, \frac{x^2}{2s}\right), \quad \text{for } x > 0, s \geq 1/2, \quad (1.1)$$

where $\Gamma(s)$ denotes the gamma function and $U(a, b, z)$ denotes the confluent hypergeometric function of the second kind (also known as the Kummer U function); see Abramowitz and Stegun (1964, Chapter 13). We denote by K_s the distribution function defined by the density κ_s . Define the Kolmogorov distance between two cumulative distribution functions P and Q as

$$d_K(P, Q) = \sup_x |P(x) - Q(x)|.$$

Theorem 1.1. *Let $W_{n,i}$ be the degree of vertex i in a preferential attachment graph on n vertices defined above and let $b_{n,i}^2 = \mathbb{E}W_{n,i}^2$. For Model 1 with*

$2 \leq i \leq n$ we have

$$d_K(\mathcal{L}(W_{n,i}/b_{n,i}), K_{i-1}) \leq \frac{C}{\sqrt{n}},$$

and for Model 2 with $1 \leq i \leq n$ we have

$$d_K(\mathcal{L}(W_{n,i}/b_{n,i}), K_{i-1/2}) \leq \frac{C}{\sqrt{n}},$$

for some constant C independent of n .

Remark 1.1. Using Proposition 2.5 below, we see an interesting difference in the behavior of the two models. In Model 1 the limit distribution for the degree of the first vertex (which by symmetry is the same as that for the second vertex) is K_1 , the absolute value of a standard normal random variable, whereas in Model 2 the limit distribution for the first vertex is $K_{1/2}$, the square root of an exponential random variable.

Remark 1.2. To ease exposition we present our results as rates, but the constants are recoverable (although probably not practical, especially for large i).

Theorem 1.1 will follow from a more general result derived by developing Stein's method for the distribution K_s . The key ingredient to our framework follows from observing that K_s is a fixed point of a certain distributional transformation which we will refer to as the “ s -transformed double size bias” (s -TDSB) transformation, which we now describe.

Recall for a non-negative random variable W , we say W' has the *size bias distribution* of W if

$$\mathbb{E}\{Wf(W)\} = \mathbb{E}W \mathbb{E}f(W'),$$

for all f such that $\mathbb{E}|Wf(W)| < \infty$; see Brown (2006) and Arratia and Goldstein (2010) for surveys and applications of size biasing. We will write W'' to denote a random variable having the size bias distribution of W' ; alternatively, we say W'' has the *double size bias distribution* of W and it is straightforward to check that

$$\mathbb{E}\{W^2f(W)\} = \mathbb{E}W^2 \mathbb{E}f(W''); \tag{1.2}$$

see also Goldstein (2007). Now, we have the following key definition.

Definition 1.3. For fixed $s \geq 1/2$, let U_1 and U_2 be two independent random variables, uniformly distributed on the interval $[0, 1]$, and let Y be a Bernoulli random variable with parameter $(2s)^{-1}$, independent of U_1 and U_2 . Define the random variable

$$V := Y \max(U_1, U_2) + (1 - Y) \min(U_1, U_2).$$

We say that W^* has the *s-transformed double size biased* (*s*-TDSB) distribution of W , if

$$\mathcal{L}(W^*) = \mathcal{L}(VW''),$$

where W'' , the double size bias of W , is assumed to be independent of V .

Our next result essentially states that the distribution K_s is the unique fixed point of the *s*-TDSB transformation. That is, $\mathcal{L}(W) = \mathcal{L}(W^*)$ if and only if $W \sim K_s$. Furthermore, the closer a distribution is to its *s*-TDSB transform, the closer it is to the K_s distribution. Besides the Kolmogorov metric, we also consider the Wasserstein metric between two probability distribution functions P and Q , defined as

$$d_W(P, Q) = \int |P(x) - Q(x)| dx.$$

Theorem 1.2. Let W be a non-negative random variable with $\mathbb{E}W^2 = 1$ and let $s \geq 1$ or $s = 1/2$. Let W^* have the *s*-TDSB distribution of W and be defined on the same probability space as W . Then, if $s \geq 1$,

$$d_W(\mathcal{L}(W), K_s) \leq 8s \left(s + \frac{1}{4} + \sqrt{\frac{\pi}{2}} \right) \mathbb{E}|W - W^*|, \quad (1.3)$$

and, for any $\beta \geq 0$,

$$d_K(\mathcal{L}(W), K_s) \leq 53s\beta + 34s^{3/2} \mathbb{P}[|W - W^*| > \beta]. \quad (1.4)$$

If $s = 1/2$, then

$$d_W(\mathcal{L}(W), K_{1/2}) \leq 2\mathbb{E}|W - W^*|.$$

and

$$d_K(\mathcal{L}(W), K_{1/2}) \leq 26\beta + 8\mathbb{P}[|W - W^*| > \beta].$$

Remark 1.4. In the case that $s = 1$, V is uniform on $(0, 1)$ so that we obtain the interesting fact that $\mathcal{L}(W) = \mathcal{L}(UW'')$ for U uniform $(0, 1)$ and independent of W'' if and only if W is the absolute value of the standard normal random variable; see Proposition 2.5 below.

Although there are abstract formulations for developing Stein's method machinery for a given distribution, see Reinert (2005), our framework below does not adhere to any of these directly since the characterizing operator we use is a second order differential operator (see (3.1) and (3.2) below). For the distribution K_s , the usual first order Stein operator derived from the density approach of Reinert (2005) (following Stein (1986)) is a complicated expression involving special functions. However, by composing this more canonical operator with an appropriate first order operator, we are able to derive a second order Stein operator (see (3.4) below) which has a form that is amenable to our analysis. This strategy may be useful for other distributions which have first order operators that are difficult to handle.

The usual approach to developing Stein's method is to decide on the distribution of interest, find a corresponding Stein operator, and then derive couplings from it. The operator we use here was suggested by the s -TDSB transform which in turn arose from the discovery of a close coupling in the preferential attachment application. We believe this approach of using couplings to suggest a Stein operator is a potentially fruitful new strategy for extending Stein's method to new distributions and applications.

There have been several previous developments of Stein's method using fixed points of distributional transformations. Goldstein and Reinert (1997) develops Stein's method using the zero-bias transformation for which the normal distribution is a fixed point. Letting U be a uniform (0,1) random variable independent of all else, Goldstein (2009) and Peköz and Röllin (2011) develop Stein's method for the exponential distribution using the fact that W and UW' have the same distribution if and only if W has an exponential distribution; Pakes and Khattree (1992) and Lyons, Pemantle, and Peres (1995) also use this property. We will show below that W and UW'' have the same distribution if and only if W is distributed as the absolute value of a standard normal random variable. In this light, this paper can be viewed as extending the use of these types of distributional transformations in Stein's method.

The layout of the remainder of the article is as follows. In Section 2 we discuss various properties and alternative representations of K_s , in Section 3 we develop Stein's method for K_s and prove Theorem 1.2, and in Section 4 we prove Theorem 1.1 by constructing the coupling needed to apply Theorem 1.2 and bounding the appropriate terms.

2 THE DISTRIBUTION K_s

In this section we collect some facts about K_s ; recall the notation and definitions associated to the formula for the density $\kappa_s(x)$. Let us collect some properties of the Kummer U function from Chapter 13 of Abramowitz and Stegun (1964). The notation $U'(a, b, z)$ refers to the derivative with respect to z , and the left italic labeling of the formulas below corresponds to the equation numbers of Abramowitz and Stegun (1964).

Lemma 2.1. *Let $a, b, z \in \mathbb{R}$.*

$$(13.1.29) \quad U(a, b, z) = z^{1-b} U(1+a-b, 2-b, z); \quad (2.1)$$

$$(13.2.5) \quad \text{if } a, z > 0, \quad U(a, b, z) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-zt} t^{a-1} (1+t)^{b-a-1} dt; \quad (2.2)$$

$$(13.4.21) \quad U'(a, b, z) = -aU(a+1, b+1, z); \quad (2.3)$$

$$(13.4.24) \quad (1+a-b)U(a, b-1, z) = (1-b)U(a, b, z) - zU'(a, b, z); \quad (2.4)$$

$$(13.4.25) \quad U(a, b, z) - U'(a, b, z) = U(a, b+1, z); \quad (2.5)$$

$$(13.4.27) \quad U(a-1, b-1, z) = (1-b+z)U(a, b, z) - zU'(a, b, z); \quad (2.6)$$

$$(13.5.2) \quad \text{for } z > 0, \quad U(a, b, z) \sim z^{-a} \quad (z \rightarrow \infty); \quad (2.7)$$

$$(13.5.10) \quad \text{for } a > -\frac{1}{2}, \quad U(a, \frac{1}{2}, 0) = \Gamma(\frac{1}{2})/\Gamma(a + \frac{1}{2}) \quad (2.8)$$

As a direct consequence of (2.5) we have

$$\frac{\partial}{\partial z}(e^{-z}U(a, b, z)) = -e^{-z}U(a, b+1, z); \quad (2.9)$$

combining (2.3) and (2.7) we find

$$U(0, b, z) = 1; \quad (2.10)$$

and using (2.1) and (2.10) implies that for $z > 0$,

$$U(-\frac{1}{2}, \frac{1}{2}, z^2) = z. \quad (2.11)$$

By comparing integrands in (2.2), we also find the following fact.

Lemma 2.2. *Let $0 < a < a'$, $b < b'$, and $z > 0$. Then*

$$\Gamma(a)U(a, b, z) > \Gamma(a')U(a', b, z) \text{ and } U(a, b, z) < U(a, b', z).$$

The next results provide simpler representations for K_s .

Proposition 2.3. *If X and Y are two independent random variables having distributions*

$$X \sim \begin{cases} B(1, s-1), & \text{if } s > 1, \\ B(1/2, s-1/2), & \text{if } 1/2 < s \leq 1, \end{cases}$$

where $B(a, b)$ denotes the beta distribution, and

$$Y \sim \begin{cases} \Gamma(1/2, 1), & \text{if } s > 1, \\ \text{Exp}(1), & \text{if } 1/2 < s \leq 1, \end{cases}$$

where $\Gamma(a, b)$ denotes the gamma distribution and $\text{Exp}(\lambda)$ the exponential distribution, then

$$\sqrt{2sXY} \sim K_s.$$

Proof. Let $s > 1$ and observe that by first conditioning on X , we can express the density of $\sqrt{2sXY}$ as

$$p_s(x) := \frac{\sqrt{2}(s-1)}{\sqrt{s\pi}} \int_0^1 \exp\left(\frac{-x^2}{2sy}\right) y^{-1/2} (1-y)^{s-2} dy. \quad (2.12)$$

After making the change of variable $y = 1/(1+t)$ in (2.12), we find

$$p_s(x) = \frac{\sqrt{2}(s-1)}{\sqrt{s\pi}} \int_0^\infty \exp\left(\frac{-x^2(t+1)}{2s}\right) t^{s-2} (1+t)^{1/2-s} dt,$$

and now using (2.2) in the definition of κ_s implies that $\kappa_s = p_s$.

Similarly, if $1/2 < s \leq 1$, then we can express the density of $\sqrt{2sXY}$ as

$$q_s(x) := \frac{\Gamma(s)x}{s\sqrt{\pi}\Gamma(s-\frac{1}{2})} \int_0^1 \exp\left(\frac{-x^2}{2sy}\right) y^{-3/2} (1-y)^{s-3/2} dy, \quad (2.13)$$

and after making the change of variable $y = 1/(1+t)$ in (2.13), we find

$$\begin{aligned} q_s(x) &= \frac{\Gamma(s)x}{s\sqrt{\pi}\Gamma(s-\frac{1}{2})} \int_0^\infty \exp\left(\frac{-x^2(t+1)}{2s}\right) t^{s-3/2} (1+t)^{1-s} dt \\ &= \Gamma(s) \sqrt{\frac{2}{s\pi}} \exp\left(\frac{-x^2}{2s}\right) \frac{x}{\sqrt{2s}} U\left(s-\frac{1}{2}, 3/2, \frac{x^2}{2s}\right), \end{aligned}$$

where we have used (2.2) in the second equality. Applying (2.1) to this last expression implies $\kappa_s = q_s$. \square

The previous representations easily yield useful formulae for Mellin transforms.

Proposition 2.4. *If $Z \sim K_s$ with $s \geq 1/2$, then for all $r > -1$,*

$$\mathbb{E}Z^r = \left(\frac{s}{2}\right)^{r/2} \frac{\Gamma(s)\Gamma(r+1)}{\Gamma(\frac{r}{2}+s)}. \quad (2.14)$$

Proof. Using Proposition 2.3 and well known formulas for the Mellin transforms of the beta and gamma distributions, we find

$$\mathbb{E}Z^r = (2s)^{r/2} \frac{\Gamma(s)\Gamma(\frac{r}{2}+1)\Gamma(\frac{r}{2}+\frac{1}{2})}{\Gamma(\frac{r}{2}+s)\Gamma(\frac{1}{2})}. \quad (2.15)$$

An application of the gamma duplication formula yields

$$\Gamma(\frac{r}{2}+1)\Gamma(\frac{r}{2}+\frac{1}{2}) = \Gamma(\frac{1}{2})2^{-r}\Gamma(r+1),$$

which combined with (2.15) implies (2.14). \square

In a few special cases we can simplify and extend Proposition 2.3. Below $K_s(x)$ denotes the distribution function of K_s .

Proposition 2.5. *We have the following special cases of K_s :*

$$\begin{aligned} (i) \quad & \kappa_{1/2}(x) = 2xe^{-x^2}, \\ (ii) \quad & \kappa_1(x) = (2/\pi)^{1/2}e^{-x^2/2}, \\ (iii) \quad & \lim_{s \rightarrow \infty} K_s(x) = 1 - e^{-\sqrt{2}x}. \end{aligned}$$

Proof. The Identities (i) and (ii) are immediate from (2.11) and (2.10), respectively. Using Stirling's formula for the gamma function to take the limit as $s \rightarrow \infty$ for fixed r in (2.14) yields the moments of $\text{Exp}(\sqrt{2})$ which proves (iii). \square

Remark 2.1. As discussed below, the preferential attachment model we study is a special case of a generalized Pólya triangular urn scheme as studied by Janson (2006). The limiting distributions in their Theorem 1.3 part (v) with $\alpha = 2$ and $\delta = \gamma = 1$ include K_s . In fact, Example 3.1 of Janson (2006) discusses these limits, but with the exception of the case $s = 1$, it does not appear that the decomposition of Proposition 2.3 has previously been exposed. On the other hand, up to a scaling factor, the moment formula of Theorem 1.7 of Janson (2006) simplifies to that of Proposition 2.4 for K_s . In a related vein, K_s also has moments of gamma type; see the survey by Janson (2010), in particular Section 9.

Additionally, if $Z \sim K_s$, then $Z^2/(2s) \sim D(1, 1/2; s)$ for $s \geq 1/2$, where $D(a, b; c)$ is a *Dufresne law* as defined in Chamayou and Letac (1999). Dufresne laws are essentially a generalization of products of independent beta and gamma random variables and they also have moments of gamma type.

We now collect one more fact about K_s , which will also prove useful in developing the Stein's method framework below.

Lemma 2.6 (Mills Ratio for K_s). *For every $x \geq 0$ and $s \geq 1$,*

$$\frac{1}{\kappa_s(x)} \int_x^\infty \kappa_s(y) dy \leq \min \left\{ \sqrt{\frac{\pi}{2}}, \frac{s}{x} \right\}.$$

Proof. By making the change of variable $\frac{y^2}{2s} = z$ and then applying (2.1) and (2.9), we find

$$\begin{aligned} \int_x^\infty \kappa_s(y) dy &= \frac{\Gamma(s)}{\sqrt{\pi}} \int_{\frac{x^2}{2s}}^\infty z^{-1/2} \exp(-z) U\left(s-1, \frac{1}{2}, z\right) dz \\ &= \frac{\Gamma(s)}{\sqrt{\pi}} \int_{\frac{x^2}{2s}}^\infty \exp(-z) U\left(s-\frac{1}{2}, \frac{3}{2}, z\right) dz \\ &= \frac{\Gamma(s)}{\sqrt{\pi}} \exp\left(-\frac{x^2}{2s}\right) U\left(s-\frac{1}{2}, \frac{1}{2}, \frac{x^2}{2s}\right), \end{aligned}$$

so that

$$\frac{1}{\kappa_s(x)} \int_x^\infty \kappa_s(y) dy = \sqrt{\frac{s}{2}} \frac{U\left(s-\frac{1}{2}, \frac{1}{2}, \frac{x^2}{2s}\right)}{U\left(s-1, \frac{1}{2}, \frac{x^2}{2s}\right)}. \quad (2.16)$$

First note that by applying (2.1) in the denominator of the final expression of (2.16) we have

$$\frac{1}{\kappa_s(x)} \int_x^\infty \kappa_s(y) dy = \frac{s}{x} \frac{U\left(s-\frac{1}{2}, \frac{1}{2}, \frac{x^2}{2s}\right)}{U\left(s-\frac{1}{2}, \frac{3}{2}, \frac{x^2}{2s}\right)} \leq \frac{s}{x}, \quad (2.17)$$

where the inequality follows by Lemma 2.2.

Now applying (2.1) in both the denominator and the numerator of (2.16), we find

$$\frac{1}{\kappa_s(x)} \int_x^\infty \kappa_s(y) dy = \sqrt{\frac{s}{2}} \frac{U\left(s, \frac{3}{2}, \frac{x^2}{2s}\right)}{U\left(s-\frac{1}{2}, \frac{3}{2}, \frac{x^2}{2s}\right)} \leq \sqrt{\frac{s}{2}} \frac{\Gamma\left(s-\frac{1}{2}\right)}{\Gamma(s)},$$

where again the inequality follows by Lemma 2.2. Now applying Lemma 2.7 below to this last expression and combining with (2.17) yields the lemma. \square

Lemma 2.7. *If $s \geq 1$, then*

$$1 \leq \frac{\sqrt{s}\Gamma(s - \frac{1}{2})}{\Gamma(s)} \leq \sqrt{\pi}$$

Proof. Theorem 1 of Bustoz and Ismail (1986) implies that

$$\frac{\sqrt{s}\Gamma(s - \frac{1}{2})}{\Gamma(s)}, \quad (2.18)$$

is a decreasing function on $(1/2, \infty)$, so that for $s \geq 1$ (2.18) is bounded above by $\sqrt{\pi}$. Moreover, Stirling's formula implies

$$\lim_{s \rightarrow \infty} \frac{\Gamma(s)}{\sqrt{s}\Gamma(s - \frac{1}{2})} = 1. \quad \square$$

3 STEIN'S METHOD FOR K_s

In this section we develop Stein's method for K_s and prove Theorem 1.2.

Lemma 3.1 (Characterizing Stein operator). *If $Z \sim K_s$ for $s \geq 1/2$, then for every twice differentiable function f with $f(0) = f'(0) = 0$ and such that $\mathbb{E}|f''(Z)|$, $\mathbb{E}|Zf'(Z)|$, and $\mathbb{E}|f(Z)|$ are finite, we have*

$$\mathbb{E}\{sf''(Z) - Zf'(Z) - 2(s-1)f(Z)\} = 0. \quad (3.1)$$

Proof. Let $C_s := \sqrt{2}\Gamma(s)/\sqrt{s\pi}$. Using integration by parts and (2.3),

$$\begin{aligned} & \mathbb{E}\{sf''(Z)\} \\ &= C_s \int_0^\infty sf''(x) \exp\left(\frac{-x^2}{2s}\right) U\left(s-1, \frac{1}{2}, \frac{x^2}{2s}\right) dx \\ &= C_s \int_0^\infty f'(x)x \exp\left(\frac{-x^2}{2s}\right) \left(U\left(s-1, \frac{1}{2}, \frac{x^2}{2s}\right) + (s-1)U\left(s, \frac{3}{2}, \frac{x^2}{2s}\right)\right) dx \\ &= \mathbb{E}\{Zf'(Z)\} + C_s \int_0^\infty f'(x) \cdot (s-1)x \exp\left(\frac{-x^2}{2s}\right) U\left(s, \frac{3}{2}, \frac{x^2}{2s}\right) dx. \end{aligned}$$

Using again integration by parts and then (2.6),

$$\begin{aligned} & C_s \int_0^\infty f'(x) \cdot x \exp\left(\frac{-x^2}{2s}\right) (s-1)U\left(s, \frac{3}{2}, \frac{x^2}{2s}\right) dx \\ &= C_s \int_0^\infty f(x) \cdot 2(s-1) \exp\left(\frac{-x^2}{2s}\right) \\ & \quad \times \left(\left(-\frac{1}{2} + \frac{x^2}{2s}\right)U\left(s, \frac{3}{2}, \frac{x^2}{2s}\right) - \frac{x^2}{2s}U'\left(s, \frac{3}{2}, \frac{x^2}{2s}\right)\right) dx \\ &= C_s \int_0^\infty f(x) \cdot 2(s-1) \exp\left(\frac{-x^2}{2s}\right) U\left(s-1, \frac{1}{2}, \frac{x^2}{2s}\right) dx \\ &= \mathbb{E}\{2(s-1)f(Z)\}. \end{aligned}$$

Hence

$$\mathbb{E}\{sf''(Z)\} = \mathbb{E}\{Zf'(Z)\} + \mathbb{E}\{2(s-1)f(Z)\}$$

which proves the claim. \square

For the sake of brevity, let $V_s(x) := U(s-1, \frac{1}{2}, \frac{x^2}{2s})$.

Lemma 3.2. *The second order differential equation*

$$sf''(x) - xf'(x) - 2(s-1)f(x) = h(x) - \mathbb{E}h(Z) \quad (3.2)$$

has solution

$$\begin{aligned} f(x) &= \frac{1}{s}V_s(x) \int_0^x \frac{1}{V_s(y)\kappa_s(y)} \int_0^y \tilde{h}(z)\kappa_s(z)dzdy \\ &= -\frac{1}{s}V_s(x) \int_0^x \frac{1}{V_s(y)\kappa_s(y)} \int_y^\infty \tilde{h}(z)\kappa_s(z)dzdy \end{aligned} \quad (3.3)$$

where $\tilde{h} = h - \mathbb{E}h(Z)$.

Proof. First we prove that we can write (3.2) as

$$sg'(x) - s(\frac{x}{s} - d(x))g(x) = \tilde{h}(x), \quad f'(x) - d(x)f(x) = g(x) \quad (3.4)$$

with

$$d(x) = \frac{\partial}{\partial x} \log V_s(x) = \frac{V'_s(x)}{V_s(x)}. \quad (3.5)$$

In order to see this, first combine the two equations of (3.4) to obtain

$$sf''(x) - xf'(x) - (sd'(x) + sd(x)^2 - xd(x))f(x) = \tilde{h}(x)$$

Hence, we only need to show that

$$sd'(x) + sd(x)^2 - xd(x) = 2(s-1). \quad (3.6)$$

In order to simplify the calculations, let us introduce

$$D(z) = \frac{\partial}{\partial z} \log U(s-1, \frac{1}{2}, z) = \frac{U'(s-1, \frac{1}{2}, z)}{U(s-1, \frac{1}{2}, z)};$$

note that $d(x) = \frac{x}{s}D(\frac{x^2}{2s})$. With this and $z = \frac{x^2}{2s}$, (3.6) becomes

$$(\frac{1}{2} - z)D(z) + zD'(z) + zD(z)^2 = s-1. \quad (3.7)$$

The left hand side of (3.7) is equal to

$$\begin{aligned} & \frac{(\frac{1}{2} - z)U'(s-1, \frac{1}{2}, z) + zU''(s-1, \frac{1}{2}, z)}{U(s-1, \frac{1}{2}, z)} \\ &= (s-1) \frac{(-\frac{1}{2} + z)U(s, \frac{3}{2}, z) - zU'(s, \frac{3}{2}, z)}{U(s-1, \frac{1}{2}, z)} = s-1, \end{aligned}$$

where we used (2.3) and then (2.6). Hence (3.6) holds and therefore (3.4) is equivalent to (3.2).

Let us now proceed to solve (3.4). Note first that the general differential equation

$$F'(x) - A'(x)F(x) = H(x), \quad x \geq 0, \quad F(0) = 0,$$

has solution

$$F(x) = e^{A(x)} \int_0^x H(z) e^{-A(z)} dz$$

Hence, noticing that

$$\frac{x}{s} - d(x) = \frac{\partial}{\partial x} \log \kappa_s(x)$$

the solution to the first equation in (3.4) is

$$g(y) = \frac{1}{\kappa_s(y)} \int_0^y \frac{\tilde{h}(z)}{s} \kappa_s(z) dz,$$

whereas the solution to the second equation in (3.4) is

$$f(x) = V_s(x) \int_0^x \frac{g(y)}{V_s(y)} dy,$$

which is the first identity of (3.3); the second follows by observing that $\int_0^\infty \tilde{h}(x) \kappa_s(x) = 0$. \square

Before developing the Stein's method machinery further we need two more lemmas, the first of which is well known.

Lemma 3.3 (Gaussian Mills Ratio). *For $x, s > 0$,*

$$\exp\left(\frac{x^2}{2s}\right) \int_x^\infty \exp\left(-\frac{t^2}{2s}\right) dt \leq \min \left\{ \sqrt{\frac{s\pi}{2}}, \frac{s}{x} \right\}.$$

Lemma 3.4. *If $d(x)$ is defined by (3.5), then for $s \geq 1$*

$$\begin{aligned} 0 &\leq -d(x) \leq \frac{\sqrt{2}\Gamma(s)}{\sqrt{s}\Gamma(s-\frac{1}{2})} < \sqrt{2} \\ 0 &\leq -xd(x) \leq 2(s-1). \end{aligned}$$

Proof. To prove the first assertion, note that (2.1) and Lemma 2.2 imply

$$\begin{aligned} -d(x) &= -\frac{x U'(s-1, \frac{1}{2}, \frac{x^2}{2s})}{s U(s-1, \frac{1}{2}, \frac{x^2}{2s})} = \frac{\sqrt{2}(s-1) U(s-1, \frac{1}{2}, \frac{x^2}{2s})}{\sqrt{s} U(s-1, \frac{1}{2}, \frac{x^2}{2s})} \\ &\leq \frac{\sqrt{2}(s-1)\Gamma(s-1)}{\sqrt{s}\Gamma(s-\frac{1}{2})}. \end{aligned}$$

The claimed bound now follows from Lemma 2.7.

For the second assertion, we use (2.4) in the second equality below to find

$$-xd(x) = -\frac{x^2 U'(s-1, \frac{1}{2}, \frac{x^2}{2s})}{s U(s-1, \frac{1}{2}, \frac{x^2}{2s})} = 2(s-\frac{1}{2}) \frac{U(s-1, -\frac{1}{2}, \frac{x^2}{2s})}{U(s-1, \frac{1}{2}, \frac{x^2}{2s})} - 1. \quad (3.8)$$

Applying Lemma 2.2 to (3.8) proves the remaining assertion. \square

Lemma 3.5. *Let g satisfy the first equation of (3.4).*

- *If h is non-negative and bounded, then for all $x > 0$ and $s \geq 1$,*

$$|g(x)| \leq \|h\| \min \left\{ \frac{1}{s} \sqrt{\frac{\pi}{2}}, \frac{1}{x} \right\}. \quad (3.9)$$

- *If h is absolutely continuous with bounded derivative, then for all $s \geq 1$*

$$\|g\| \leq \|h'\| \left(1 + \frac{1}{s} \sqrt{\frac{\pi}{2}} \right). \quad (3.10)$$

Proof. It is easy to verify that we may write

$$g(x) = \frac{1}{s\kappa_s(x)} \int_0^x \kappa_s(y) \tilde{h}(y) dy = -\frac{1}{s\kappa_s(x)} \int_x^\infty \kappa_s(y) \tilde{h}(y) dy.$$

Thus if $h > 0$ with $\|h\| < \infty$, then for all $s \geq 1$,

$$|g(x)| \leq \frac{\|\tilde{h}\|}{s\kappa_s(x)} \int_x^\infty \kappa_s(y) dy \leq \min \left\{ \sqrt{\frac{\pi}{2}}, \frac{s}{x} \right\} \frac{\|h\|}{s},$$

where we have used Lemma 2.6; this shows (3.9).

If $\|h'\| < \infty$, then without loss of generality we can assume that $h(0) = 0$ so that for $x \geq 0$, $|h(x)| \leq \|h'\|x$. In particular, $\tilde{h}(x) \leq (x + \mathbb{E}Z_s)\|h'\|$ so noting that $\mathbb{E}Z_s \leq \sqrt{\mathbb{E}Z_s^2} = 1$, we can now use Lemma 2.6 to find

$$|g(x)| \leq \frac{\|h'\|}{s\kappa_s(x)} \int_x^\infty (y+1)\kappa_s(y) dy \leq \frac{\|h'\|}{s} \left(\frac{\int_x^\infty y\kappa_s(y) dy}{\kappa_s(x)} + \sqrt{\frac{\pi}{2}} \right).$$

To bound the integral in this last expression, we make the change of variable $\frac{y^2}{2s} = z$ and apply (2.9) to find

$$\frac{\int_x^\infty y \kappa_s(y) dy}{\kappa_s(x)} = \frac{s}{\kappa_s(x)} \int_{\frac{x^2}{2s}}^\infty e^{-z} U(s-1, \frac{1}{2}, z) dz = s \frac{U(s-1, -\frac{1}{2}, \frac{x^2}{2s})}{U(s-1, \frac{1}{2}, \frac{x^2}{2s})} \leq s,$$

where the last inequality follows from Lemma 2.2. \square

Lemma 3.6. *Let f be defined as in (3.3).*

- *If h is non-negative and bounded and $s \geq 1$, then*

$$\|f'\| \leq \sqrt{2\pi} \|h\|. \quad (3.11)$$

- *If h is non-negative, bounded and absolutely continuous with bounded derivative and $s \geq 1$, then*

$$\|f''\| \leq 2 \left(\pi\sqrt{s} + \frac{1}{s} \right) \|h\|. \quad (3.12)$$

If $s = 1/2$, then

$$\|f''\| \leq 4 \|h\|. \quad (3.13)$$

- *If h is absolutely continuous with bounded derivative and $s \geq 1$, then*

$$\|f'''\| \leq 8 \left(s + \frac{1}{4} + \sqrt{\frac{\pi}{2}} \right) \|h'\|. \quad (3.14)$$

If $s = 1/2$, then

$$\|f'''\| \leq 4 \|h'\|. \quad (3.15)$$

Proof. If either h is bounded or absolutely continuous with bounded derivative, then Lemma 3.5 implies that f satisfies the second equation of (3.4) for g bounded. Thus

$$f(x) = V_s(x) \int_0^x \frac{g(y)}{V_s(y)} dy.$$

If $s \geq 1$, then since $V_s(x) = U(s-1, \frac{1}{2}, \frac{x^2}{2s})$ is non-increasing for positive x , we find

$$|f(x)| \leq x \|g\|. \quad (3.16)$$

Now, again by (3.4), we have

$$|f'(x)| \leq |d(x)f(x)| + \|g\| \leq \|g\| (|xd(x)| + 1) \leq \|g\| (2s-1), \quad (3.17)$$

where we have used (3.16) in the first inequality and Lemma 3.4 in the second. Applying the bound (3.9) proves (3.11).

To bound f'' for h having $\|h'\| < \infty$, let $s \geq 1/2$ and differentiate (3.2) to find

$$f'''(x) - \frac{x}{s}f''(x) = \frac{2s-1}{s}f'(x) + \frac{h'(x)}{s}$$

which implies

$$\frac{d}{dx} \left(\exp\left(\frac{-x^2}{2s}\right) f''(x) \right) = \exp\left(\frac{-x^2}{2s}\right) \left(\frac{2s-1}{s} f'(x) + \frac{h'(x)}{s} \right).$$

Integrating, we obtain

$$\exp\left(\frac{-x^2}{2s}\right) f''(x) = - \int_x^\infty \exp\left(\frac{-y^2}{2s}\right) \left(\frac{2s-1}{s} f'(y) + \frac{h'(y)}{s} \right) dy,$$

so that Lemma 3.3 yields

$$|f''(x)| \leq (2s-1)\|f'\| \min \left\{ \sqrt{\frac{\pi}{2s}}, \frac{1}{x} \right\} + \frac{1}{s} \exp\left(\frac{-x^2}{2s}\right) \int_x^\infty \exp\left(\frac{-y^2}{2s}\right) h'(y) dy. \quad (3.18)$$

If $\|h\| < \infty$, then an integration by parts yields a bound on the second term of (3.18) which yields

$$|f''(x)| \leq (2s-1)\|f'\| \min \left\{ \sqrt{\frac{\pi}{2s}}, \frac{1}{x} \right\} + \frac{2\|h\|}{s}.$$

If $s \geq 1$, then apply the bound (3.11) above on $\|f'\|$ to find (3.12); for $s = 1/2$, (3.13) follows immediately. Alternatively, we can apply Lemma 3.3 directly to (3.18) to find

$$|f''(x)| \leq ((2s-1)\|f'\| + \|h'\|) \min \left\{ \sqrt{\frac{\pi}{2s}}, \frac{1}{x} \right\}. \quad (3.19)$$

Finally, differentiating (3.2) yields

$$s|f'''(x)| \leq |xf''(x)| + (2s-1)\|f'\| + \|h'\|; \quad (3.20)$$

the first term can be bounded by (3.19), and if $s \geq 1$ a subsequent application of (3.17) on $\|f'\|$ and then (3.10) on $\|g\|$ yields (3.14). If $s = 1/2$, then (3.15) follows from (3.20) and (3.19). \square

In order to obtain the bounds for the Kolmogorov metric, we need to introduce the smoothed half-line indicator function

$$h_{a,\varepsilon}(x) = \frac{1}{\varepsilon} \int_0^\varepsilon \mathbb{I}[x \leq a + s] ds. \quad (3.21)$$

Lemma 3.7. *If $Z \sim K_s$ and W is a non-negative random variable and $s \geq 1$, then, for all $\varepsilon > 0$,*

$$d_K(\mathcal{L}(W), K_s) \leq \sup_{a \geq 0} |\mathbb{E}h_{a,\varepsilon}(W) - \mathbb{E}h_{a,\varepsilon}(Z)| + \varepsilon\sqrt{2}.$$

If $s = 1/2$, then, for all $\varepsilon > 0$,

$$d_K(\mathcal{L}(W), K_{1/2}) \leq \sup_{a \geq 0} |\mathbb{E}h_{a,\varepsilon}(W) - \mathbb{E}h_{a,\varepsilon}(Z)| + \varepsilon\sqrt{2/e}.$$

Proof. The lemma follows from a well known argument and the following bounds on the density $\kappa_s(x)$. If $s \geq 1$, then $\kappa_s(x)$ is decreasing in x and from (2.8),

$$\kappa_s(0) = \frac{\Gamma(s)\sqrt{2}}{\Gamma(s - \frac{1}{2})\sqrt{s}} \leq \sqrt{2},$$

where the inequality is by Lemma 2.7. If $s = 1/2$, then $\kappa_s(x) = 2xe^{-x^2}$ which has maximum $\sqrt{2/e}$. \square

We will also need the following “indirect” concentration inequality; it follows from the arguments of the proof of Lemma 3.7 immediately above.

Lemma 3.8. *If $Z \sim K_s$ and W is a non-negative random variable and $s \geq 1$, then, for all $0 \leq a < b$,*

$$\mathbb{P}(a < W \leq b) \leq \sqrt{2}(b - a) + 2d_K(\mathcal{L}(W), K_s).$$

If $s = 1/2$, then, for all $0 \leq a < b$,

$$\mathbb{P}(a < W \leq b) \leq \sqrt{2/e}(b - a) + 2d_K(\mathcal{L}(W), K_s).$$

Lemma 3.9. *If f satisfies (3.2) for $h_{a,\varepsilon}$ and $s \geq 1$, then*

$$\begin{aligned} s|f''(x+t) - f''(x)| &\leq |t| \left(2x(\pi\sqrt{s} + 1) + (2s-1)\sqrt{2\pi} \right) \\ &\quad + \frac{1}{\varepsilon} \int_{t \wedge 0}^{t \vee 0} \mathbb{I}[a < x + u \leq a + \varepsilon] du. \end{aligned}$$

If $s = 1/2$, then

$$\frac{1}{2}|f''(x+t) - f''(x)| \leq 4|t|x + \frac{1}{\varepsilon} \int_{t \wedge 0}^{t \vee 0} \mathbb{I}[a < x + u \leq a + \varepsilon] du.$$

Proof. Using (3.2), we obtain

$$\begin{aligned} s(f''(x+t) - f''(x)) &= x(f'(x+t) - f'(x)) + tf'(x+t) \\ &\quad + 2(s-1)(f(x+t) - f(x)) + h_{a,\varepsilon}(x+t) - h_{a,\varepsilon}(x), \end{aligned}$$

hence

$$\begin{aligned} s|f''(x+t) - f''(x)| &\leq |t| (x\|f''\| + \|f'\| + 2(s-1)\|f'\|) \\ &\quad + \frac{1}{\varepsilon} \int_{t \wedge 0}^{t \vee 0} \mathbf{I}[a < x+u \leq a+\varepsilon] du. \end{aligned}$$

Applying the bounds of Lemma 3.6 yields the claim. \square

Lemma 3.10. *Let W be a non-negative random variable with $\mathbb{E}W^2 = 1$ and let W^* be the s -TDSB of W as in Definition 1.3 for some $s \geq 1/2$. For every twice differentiable function f with $f(0) = f'(0) = 0$ and such that the expectations below are well defined, we have*

$$s\mathbb{E}f''(W^*) = \mathbb{E}\{Wf'(W) + 2(s-1)f(W)\}$$

Proof. The lemma will follow from two facts:

- If W'' has the double size bias distribution of W , then for all g with $\mathbb{E}|W^2g(W)| < \infty$,

$$\mathbb{E}g(W'') = \mathbb{E}\{W^2g(W)\}.$$

- If g is a function such that $g'(0) = g(0) = 0$ with $\mathbb{E}|g''(V)| < \infty$, then

$$s\mathbb{E}g''(V) = g'(1) + 2(s-1)g(1).$$

The first item above is easy to verify from the definition of the size bias distribution and the fact that $\mathbb{E}W^2 = 1$, and the second follows from a simple calculation after noting that V has density $(2 - \frac{1}{s}) - 2x(1 - \frac{1}{s})$ for $0 < x < 1$.

By conditioning on W'' and using the second fact above for $g(t) = f(tW'')/(W'')^2$, we find

$$s\mathbb{E}f''(W^*) = \mathbb{E}\left\{\frac{f'(W'')}{W''} + 2(s-1)\frac{f(W'')}{(W'')^2}\right\},$$

and applying the first fact above proves the lemma. \square

Proof of Wasserstein bound of Theorem 1.2. Making use of Lemma 3.2 and Lemma 3.10, we have

$$\begin{aligned}\mathbb{E}h(W) - \mathbb{E}h(Z) &= \mathbb{E}\{sf''(W) - Wf'(W) - 2(s-1)f(W)\} \\ &= s\mathbb{E}\{f''(W) - f''(W^*)\},\end{aligned}$$

where f is given by (3.3). If h is Lipschitz continuous, then f is three times differentiable almost everywhere and we have

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq s\|f'''\|\mathbb{E}|W - W^*|.$$

We now obtain (1.3) by invoking (3.14) and (3.15) of Lemma 3.6. \square

Proof of Kolmogorov bound of Theorem 1.2. Fix $a > 0$ and let $\varepsilon > 0$, to be chosen later. Let f as in (3.3) with respect to $h_{a,\varepsilon}$ from (3.21). Define the indicator random variable $J = \mathbb{I}[|W - W^*| \leq \beta]$. Now,

$$\begin{aligned}\mathbb{E}h_{a,\varepsilon}(W) - \mathbb{E}h_{a,\varepsilon}(Z) &= s\mathbb{E}\{f''(W) - f''(W^*)\} \\ &= s\mathbb{E}\{J(f''(W) - f''(W^*))\} + s\mathbb{E}\{(1-J)(f''(W) - f''(W^*))\} \\ &=: R_1 + R_2.\end{aligned}$$

If $s \geq 1$, using (3.12) from Lemma 3.6 implies

$$|R_2| \leq 4(\pi s^{3/2} + 1)\mathbb{P}(|W - W^*| > \beta) \leq 17s^{3/2}\mathbb{P}(|W - W^*| > \beta).$$

Applying Lemma 3.9,

$$|R_1| \leq \beta(2\mathbb{E}W(\pi\sqrt{s} + 1) + (2s-1)\sqrt{2\pi}) + \frac{1}{\varepsilon} \int_{-\beta}^{\beta} \mathbb{P}(a < W + u \leq a + \varepsilon) du.$$

Noticing that $\mathbb{E}W \leq 1$ and applying Lemma 3.8 to the integrand,

$$|R_1| \leq 12s\beta + 2\beta\varepsilon^{-1}(\sqrt{2\varepsilon} + 2\delta) \leq 15s\beta + 4\beta\varepsilon^{-1}\delta$$

where $\delta = d_K(\mathcal{L}(W), K_s)$.

From Lemma 3.7, we have

$$\delta \leq \sqrt{2\varepsilon} + 15s\beta + 4\beta\varepsilon^{-1}\delta + 17s^{3/2}\mathbb{P}(|W - W^*| > \beta).$$

Choosing $\varepsilon = 8\beta$ and solving for δ ,

$$\delta \leq 16\sqrt{2}s\beta + 30s\beta + 34s^{3/2}\mathbb{P}(|W - W^*| > \beta).$$

which yields (1.4).

A nearly identical argument yields the statement for $s = 1/2$. \square

4 PROOF OF THEOREM 1.1

We first reformulate Theorem 1.1 in terms of a generalized Pólya urn model. An urn initially contains i black balls and j white balls and at each step a ball is drawn. If the ball drawn is black, it is returned to the urn along with an additional α black balls and β white balls; if the ball drawn is white, the ball is returned to the urn along with an additional γ black balls and δ white balls. We use the notation $(\alpha, \beta; \gamma, \delta)_{i,j}^n$ to denote the distribution of the number of white balls in this model after n draws and replacements. For example, $(\alpha, \beta; \gamma, \delta)_{i,j}^0$ has a single point mass at j .

Theorem 1.1b. *Let $n \geq 1$ and $i \geq 0$ be integers and $\mathcal{L}(W_{n,i}) = (2, 0; 1, 1)_{i,1}^n$. If $b_{n,i}^2 = \mathbb{E}W_{n,i}^2$, then, for some constant C independent of n ,*

$$d_K(\mathcal{L}(W_{n,i}/b_{n,i}), K_{(i+1)/2}) \leq \frac{C}{\sqrt{n}}.$$

Theorem 1.1 follows immediately from Theorem 1.1b after noting that for Model 1 with $i \geq 2$, the degree of the i th vertex when the graph has n vertices has distribution $(2, 0; 1, 1)_{2i-3,1}^{n-i}$, and for Model 2 with $i \geq 1$, the degree of the i th vertex when the graph has n vertices has distribution $(2, 0; 1, 1)_{2i-2,1}^{n-i+1}$.

Let $W := W_{n,i}$ have distribution $(2, 0; 1, 1)_{i,1}^n$. We will use (1.4) to prove Theorem 1.1b and so we will show that there is a close coupling of W and VW'' , where V and Y are as in Definition 1.3 with $s = (i+1)/2$. This result will follow from the following lemmas proved at the end of this section.

Lemma 4.1. *There is a coupling (Z, W'') of $(2, 0; 1, 1)_{i,3}^{n-1}$ and the double size bias distribution of $(2, 0; 1, 1)_{i,1}^n$ satisfying $\mathbb{P}(Z \neq W'') \leq C/\sqrt{n}$.*

Lemma 4.2.

- *The distribution of $(2, 0; 1, 1)_{i,1}^n$ given the first ball drawn is white is equal to $(2, 0; 1, 1)_{i+1,2}^{n-1}$.*
- *The distribution of $(2, 0; 1, 1)_{i,1}^n$ given the first ball drawn is black is equal to $(2, 0; 1, 1)_{i+2,1}^{n-1}$.*

Lemma 4.3. *For Z as in Lemma 4.1,*

$$(2, 0; 1, 1)_{i+1,2}^{n-1} = (1, 0; 0, 1)_{1,2}^{Z-3} \quad \text{and} \quad (2, 0; 1, 1)_{i+2,1}^{n-1} = (1, 0; 0, 1)_{2,1}^{Z-3},$$

where the notation $(1, 0; 0, 1)_{1,2}^{Z-3}$ is understood by the relation $[(1, 0; 0, 1)_{1,2}^{Z-3} | Z = z] = (1, 0; 0, 1)_{1,2}^{z-3}$.

Lemma 4.4. *Let U_1 and U_2 uniform $(0, 1)$ random variables, independent of each other and of Z , defined as in Lemma 4.1. Then there exists random variables X_1 with distribution $(1, 0; 0, 1)_{1,2}^{Z-3}$ and X_2 with distribution $(1, 0; 0, 1)_{2,1}^{Z-3}$ such that*

$$|X_1 - Z \max(U_1, U_2)| < 3 \quad \text{and} \quad |X_2 - Z \min(U_1, U_2)| < 3.$$

From these lemmas, we can now easily prove Theorem 1.1b; here and below we use C to denote a generic constant that may differ from line to line.

Proof of Theorem 1.1b. Let (Z, W'') be defined as in 4.1 above and let Y be a Bernoulli($1/(1+i)$) random variable; the parameter is equal to the probability that the first ball drawn under the $(2, 0; 1, 1)_{i,1}$ urn rule is white. Lemmas 4.2, 4.3, and 4.4 imply that we can couple W and VZ together so that $|W - VZ| < 3$ almost surely. Thus, using Lemma 4.1,

$$\mathbb{P}(|W - VW''| > 3) \leq \mathbb{P}(W'' \neq Z) \leq C/\sqrt{n},$$

and the theorem follows from (1.4) taking $\beta = 3/b$, noting that for $c > 0$, $(cW)' \stackrel{\mathcal{D}}{=} cW'$, and using $b^2 \geq Cn$ from Lemma 4.7. \square

We have left to prove Lemmas 4.1, 4.2, 4.3, and 4.4. The proof of Lemma 4.1 is the most involved, so we postpone it until the end of this section. Lemma 4.2 is obvious and the proofs for Lemmas 4.3 and 4.4 immediately follow.

Proof of Lemma 4.3. Consider an urn with i green balls, 1 black ball and 2 white balls. A ball is drawn at random and replaced in the urn along with another ball of the same color plus an additional green ball.

If X is the number of times a non-green ball is drawn in $n-1$ draws, the number of white balls after $n-1$ draws is distributed as $(1, 0; 0, 1)_{1,2}^X$. Since $X+3$ is distributed as $(2, 0; 1, 1)_{i,3}^{n-1}$ and the number of white balls after $n-1$ draws has distribution $(2, 0; 1, 1)_{i+1,2}^{n-1}$, the first equation follows. The second equation follows from similar considerations. \square

Proof of Lemma 4.4. We will show that for U_1 and U_2 independent uniform $(0, 1)$ random variables, there exists random variables N and M such that $\mathcal{L}(N) = (1, 0; 0, 1)_{1,2}^{n-3}$, $\mathcal{L}(M) = \mathcal{L}(n-N)$, and

$$|N - n \max(U_1, U_2)| < 3 \quad \text{and} \quad |M - n \min(U_1, U_2)| < 3 \text{ a.s.}$$

The formulas of Durrett (2010, p. 206) imply that $(1, 0; 0, 1)_{1,2}^{n-3}$ has distribution function

$$F(k) = \left(\frac{k}{n-1} \right) \left(\frac{k-1}{n-2} \right), \quad k = 1, \dots, n-1, \quad (4.1)$$

and it is straightforward to verify that

$$N := \max(\lceil (n-1)U_1 \rceil, 1 + \lceil (n-2)U_2 \rceil),$$

has the same distribution. We find $|N - n \max(U_1, U_2)| < 3$ and thus a coupling satisfying the first claim above. Defining

$$M := \min(\lceil (n-1)U_1 \rceil, 1 + \lceil (n-2)U_2 \rceil),$$

(4.1) implies $\mathcal{L}(M) = \mathcal{L}(n - N)$, and $|M - n \min(U_1, U_2)| < 3$. \square

We now only need to prove Lemma 4.1, which is the content of the remainder of this section. The next lemma helps provide a useful construction for the double size bias distribution of a sum of indicators.

Lemma 4.5. *Let $W = \sum_{i=1}^n X_i$, where the X_i are Bernoulli random variables and $b^2 := \mathbb{E}W^2$. For each $j, k \in \{1, \dots, n\}$, let $(X_i^{(j,k)})_{i \neq j,k}$ have the distribution of $(X_i)_{i \neq j,k}$ conditional on $X_j = X_k = 1$ and let J and K be random variables independent of the variables above satisfying*

$$\mathbb{P}(J = j, K = k) = \frac{\mathbb{E}(X_j X_k)}{b^2}, \quad j, k \geq 1.$$

Note that J and K can be equal and in this case the notation $i \neq j, k$ reduces to $i \neq j$. In this notation,

$$W'' = \sum_{i \neq J, K} X_i^{(J, K)} + 2 - \mathbb{I}[J = K]$$

has the double size bias distribution of W .

Proof. Note that

$$\begin{aligned} \mathbb{E}f(W'') &= b^{-2} \sum_{j,k} \mathbb{E}(X_j X_k) \mathbb{E}\{f(W) | X_j = X_k = 1\} \\ &= b^{-2} \sum_{j,k} \mathbb{E}\{X_j X_k f(W)\} = b^{-2} \mathbb{E}\{W^2 f(W)\}; \end{aligned}$$

this is exactly (1.2). \square

To simplify the notation we consider i fixed in what follows. We write

$$W_n = \sum_{j=0}^n X_j,$$

where for $j \geq 1$, X_j is the indicator that a white ball is drawn on draw j from the $(2, 0; 1, 1)_{i,1}$ urn and $X_0 = 1$ to represent the initial white ball in the urn. We will then define random variables $M_n^{j,k}$ such that

$$\mathcal{L}(M_n^{j,k}) = \mathcal{L}(W | X_j = X_k = 1), \quad (4.2)$$

so that by Lemma 4.5, if J and K are random variables independent of $M_n^{j,k}$ satisfying

$$\mathbb{P}(K = k, J = j) = \frac{\mathbb{E}(X_k X_j)}{b^2}, \quad j, k \geq 0, \quad (4.3)$$

for $b^2 := \mathbb{E}W^2$, then $M_n^{J,K}$ has the double size bias distribution of W .

In order to generate a variable satisfying (4.2) for $j < k$, we use the following lemma that yields a method to construct an urn process having the law of the $(2, 0; 1, 1)_{i,1}$ urn process up to time n conditional on $X_k = X_j = 1$. This conditioned process follows the law of the $(2, 0, 1, 1)_{i,3}$ urn process up to (and including) draw $j - 1$. At draw j , exactly one black ball is added and then draws $j + 1$ through $k - 1$ follow the $(2, 0; 1, 1)$ urn law. Again at draw k exactly one black ball is added and then the process continues to draw n following the $(2, 0; 1, 1)$ urn rule. We write $M_n^{j,k}$ to denote the number of white balls after n draws for this process, and we refer to this as the $M^{j,k}$ process. The next lemma shows that this construction of $M_n^{j,k}$ has the distribution specified in (4.2).

Lemma 4.6. *Retaining the notation and definitions above and letting $A_m = \{X_m = 1\}$, if $1 \leq j < k \leq n$ then*

$$\mathbb{P}(A_j | A_k, W_{j-1}) = \frac{1 + W_{j-1}}{2j + i},$$

and if $1 \leq l < j < k \leq n$

$$\mathbb{P}(A_l | A_k, A_j, W_{l-1}) = \frac{2 + W_{l-1}}{2l + i + 1}.$$

Proof. By Bayes' rule we have

$$\mathbb{P}(A_l | A_k, A_j, W_{l-1}) = \frac{\mathbb{P}(A_l | W_{l-1}) \mathbb{P}(A_k A_j | A_l, W_{l-1})}{\mathbb{P}(A_k A_j | W_{l-1})} \quad (4.4)$$

and

$$\mathbb{P}(A_j|A_k, W_{j-1}) = \frac{\mathbb{P}(A_j|W_{j-1})\mathbb{P}(A_k|A_j, W_{j-1})}{\mathbb{P}(A_k|W_{j-1})} \quad (4.5)$$

and we next will calculate the probabilities above. For $j \geq 1$, we have

$$\mathbb{P}(A_j|W_{j-1}) = \frac{W_{j-1}}{i + 2j - 1} \quad (4.6)$$

which implies that for $k \geq j$,

$$\mathbb{P}(A_k|W_{j-1}) = \frac{\mathbb{E}(W_{k-1}|W_{j-1})}{i + 2k - 1}$$

and

$$\mathbb{P}(A_k|A_j, W_{j-1}) = \frac{\mathbb{E}(W_{k-1}|A_j, W_{j-1})}{i + 2k - 1}. \quad (4.7)$$

Now, to compute the conditional expectations appearing above, note first that

$$\mathbb{E}(W_k|W_{k-1}) = W_{k-1} + \frac{W_{k-1}}{i + 2k - 1} = \left(\frac{i + 2k}{i + 2k - 1} \right) W_{k-1}, \quad (4.8)$$

and conditioning on W_{j-1} and taking expectations yields

$$\mathbb{E}(W_k|W_{j-1}) = \left(\frac{i + 2k}{i + 2k - 1} \right) \mathbb{E}(W_{k-1}|W_{j-1})$$

and then iterating and substituting $k - 1$ for k yields

$$\mathbb{E}(W_{k-1}|W_{j-1}) = \prod_{t=1}^{k-j} \left(\frac{i + 2(k-t)}{i + 2(k-t) - 1} \right) W_{j-1} \quad (4.9)$$

and also

$$\mathbb{E}(W_{k-1}|A_j, W_{j-1}) = \prod_{t=1}^{k-j-1} \left(\frac{i + 2(k-t)}{i + 2(k-t) - 1} \right) (1 + W_{j-1}). \quad (4.10)$$

We use a similar approach to obtain

$$\begin{aligned} \mathbb{E}(W_k^2|W_{k-1}) &= W_{k-1}^2 \left(1 - \frac{W_{k-1}}{i + 2k - 1} \right) + (W_{k-1} + 1)^2 \frac{W_{k-1}}{i + 2k - 1} \\ &= \left(\frac{i + 2k + 1}{i + 2k - 1} \right) W_{k-1}^2 + \frac{W_{k-1}}{i + 2k - 1}. \end{aligned}$$

which can then be added to (4.8) while letting $D_k = W_k(1 + W_k)$ to obtain

$$\mathbb{E}(D_k|W_{k-1}) = \frac{i + 2k + 1}{i + 2k - 1} D_{k-1}$$

and thus

$$\mathbb{E}(D_k|W_{j-1}) = \frac{i + 2k + 1}{i + 2k - 1} \mathbb{E}(D_{k-1}|W_{j-1}).$$

Iterating and substituting $k - 1$ for k gives

$$\mathbb{E}(D_{k-1}|W_{j-1}) = \frac{i + 2k - 1}{i + 2j - 1} D_{j-1} = \frac{i + 2k - 1}{i + 2j - 1} W_{j-1}(1 + W_{j-1}) \quad (4.11)$$

and also

$$\mathbb{E}(D_{k-1}|A_j, W_{j-1}) = \frac{i + 2k - 1}{i + 2j + 1} (W_{j-1} + 1)(W_{j-1} + 2).$$

Letting

$$c = \frac{1}{(i + 2j - 1)(i + 2k - 1)} \prod_{t=1}^{k-j-1} \left(\frac{i + 2(k - t)}{i + 2(k - t) - 1} \right)$$

and applying (4.7) and (4.10) we have

$$\begin{aligned} \mathbb{P}(A_j A_k | W_{l-1}) &= \mathbb{E}(\mathbb{P}(A_j | W_{j-1}) \mathbb{P}(A_k | A_j, W_{j-1}) | W_{l-1}) \\ &= c \mathbb{E}(D_{j-1} | W_{l-1}) \\ &= c \frac{i + 2j - 1}{i + 2l - 1} W_{l-1}(1 + W_{l-1}) \end{aligned}$$

which also implies

$$\mathbb{P}(A_j A_k | A_l, W_{l-1}) = c \frac{i + 2j - 1}{i + 2l + 1} (1 + W_{l-1})(2 + W_{l-1}).$$

Substituting these expressions appropriately into (4.4) and (4.5) proves the lemma. \square

We are now ready to prove Lemma 4.1.

Proof of Lemma 4.1. The $(2, 0; 1, 1)_{i,3}$ process and the $M^{j,k}$ process defined above differ only in that in the latter process, after either draw j or k a single black ball is added into the urn regardless of what is drawn; in the former process, two balls are always added to the urn depending on the

color drawn. This difference turns out to be small enough to allow a close coupling as stated in the lemma.

For each t we will construct $(z_t, m_t^{j,k})$ to respectively denote the indicator for the event that a white ball is added to the urn after draw number t for the $(2, 0; 1, 1)_{i,3}$ process and for the $M^{j,k}$ process, and we write

$$Z_n = \sum_{t=0}^n z_t, \quad M_n^{j,k} = \sum_{t=0}^n m_t^{j,k}$$

to denote the number of white balls in the urn after draw n for each process. Let U_t be independent uniform $(0, 1)$ random variables. We define

$$z_t = \mathbb{I}\left[U_t < \frac{Z_{t-1}}{i + 2t + 1}\right]$$

and for $t \neq k, t \neq j$ we define

$$m_t^{j,k} = \mathbb{I}\left[U_t < \frac{M_{t-1}^{j,k}}{i + 2t + 1 - \mathbb{I}\{t > j\} - \mathbb{I}\{t > k\}}\right].$$

We also set $m_k^{j,k} = m_j^{j,k} = 0$ since a single black ball is added after draws j and k . Using that the event $M_n^{j,k} \neq Z_{n-1}$ can be written as a union of the events that index t is the least index such that $z_t \neq m_t^{j,k}$, we have for $0 < j < k$,

$$\begin{aligned} & \mathbb{P}(M_n^{j,k} \neq Z_{n-1}) \\ & \leq \mathbb{E}(z_k + z_j + m_n) + \sum_{t=j}^n \mathbb{P}\left(\frac{Z_{t-1}}{i + 2t + 1} < U_t < \frac{Z_{t-1}}{i + 2t - 1}\right) \\ & \leq 9\mathbb{E}X_j + 3 \sum_{t=j}^n \mathbb{E}W_{t-1} \left(\frac{1}{i + 2t - 1} - \frac{1}{i + 2t + 1}\right) \\ & \leq C\sqrt{\frac{j}{i}} \left(\frac{1}{i + 2j + 1}\right) + C \sum_{t=j}^n \sqrt{\frac{t}{i}} \left(\frac{1}{i + 2t - 1} - \frac{1}{i + 2t + 1}\right) \\ & \leq Cj^{-\frac{1}{2}}, \end{aligned}$$

where we have used Lemma 4.7 below and $\mathbb{E}z_j \leq 3\mathbb{E}X_j$. Defining J, K as

in (4.3) we then have

$$\begin{aligned}
& \mathbb{P}(M_n^{J,K} \neq Z_{n-1}) \\
& \leq \mathbb{P}(J = 0) + \mathbb{P}(K = 0) + \mathbb{P}(J = K) + \frac{2C}{b^2} \sum_{j < k} j^{-\frac{1}{2}} \mathbb{E}X_j X_k \\
& \leq (2 + \mathbb{E}W_n)/b^2 + \frac{C}{b^2} \sum_{j < k} \frac{1}{\sqrt{jk}} \frac{1}{\sqrt{j}} \\
& \leq C/\sqrt{n} + C \sum_j j^{-3/2} \leq C/\sqrt{n}. \quad \square
\end{aligned}$$

Lemma 4.7. Fix $i \geq 1$ and let $W_n = \sum_{j=0}^n X_j$ where for $j \geq 1$, X_j is the indicator that a white ball is drawn on draw j from the $(2, 0; 1, 1)_{i,1}$ urn and $X_0 = 1$. Retaining the other notation and definitions above, for all $1 \leq j < k \leq n$,

$$\begin{aligned}
\mathbb{E}W_n & \leq \sqrt{2\pi} \sqrt{\frac{n}{i+2} + \frac{1}{2}}, \quad \mathbb{E}W_n^2 \geq (2 - \sqrt{\pi}) \frac{i + 2n + 1}{i + 1}, \\
\mathbb{E}(X_j X_k) & \leq \frac{\sqrt{2\pi}(1 + \sqrt{\pi})}{(i+2)\sqrt{(i+2j)(i+2k-1)}}, \\
\mathbb{E}X_j & \leq \frac{\sqrt{\pi}}{\sqrt{(i+1)(i+2j-1)}}.
\end{aligned}$$

Proof. From (4.9) we have

$$\mathbb{E}W_n = \prod_{t=1}^n \frac{i+2t}{i+2t-1} = \frac{\Gamma(\frac{i+1}{2}) \Gamma(n + \frac{i+1}{2} + \frac{1}{2})}{\Gamma(\frac{i+1}{2} + \frac{1}{2}) \Gamma(n + \frac{i+1}{2})}, \quad (4.12)$$

from (4.6) we find

$$\mathbb{E}X_j = \frac{\mathbb{E}W_{j-1}}{i+2j-1},$$

and using (4.7) and (4.10) yields

$$\mathbb{E}(X_j X_k) = \frac{\mathbb{E}(1 + W_{j-1})\mathbb{E}X_j}{i+2k-1} \prod_{t=j+1}^{k-1} \frac{i+2t}{i+2t-1}.$$

Now using (4.11) we find

$$\mathbb{E}W_n^2 = 2 \frac{i+2n+1}{i+1} - \mathbb{E}W_n.$$

Lemma 2.7 applied to (4.12) implies

$$\frac{1}{\sqrt{\pi}}\sqrt{\frac{2n}{i+2}}+1 \leq \prod_{t=1}^n \frac{i+2t}{i+2t-1} \leq \sqrt{\pi}\sqrt{\frac{2n}{i+2}}+1,$$

and collecting the facts above yields the lemma. \square

ACKNOWLEDGMENTS

The authors would like to thank Carl Morris for valuable advice on this project.

The authors have learned that Robert Gaunt (University of Oxford) has independently been working on a second order Stein operator for the variance-gamma distribution, but there appears to be little overlap between their work and ours.

REFERENCES

- M. Abramowitz and I. A. Stegun, editors (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover Publications Inc., New York.
- R. Arratia and L. Goldstein (2010). Size bias, sampling, the waiting time paradox, and infinite divisibility: when is the increment independent? *Preprint*. Available at <http://arxiv.org/abs/1007.3910>.
- Á. Backhausz (2011). Limit distribution of degrees in random family trees. *Electron. Commun. Probab.* **16**, 29–37. ISSN 1083-589X.
- A.-L. Barabási and R. Albert (1999). Emergence of scaling in random networks. *Science* **286**, 509–512.
- B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády (2001). The degree sequence of a scale-free random graph process. *Random Structures Algorithms* **18**, 279–290.
- M. Brown (2006). Exploiting the waiting time paradox: applications of the size-biasing transformation. *Probability in the Engineering and Informational Sciences* **20**, 195–230.
- J. Bustoz and M. E. H. Ismail (1986). On gamma function inequalities. *Math. Comp.* **47**, 659–667.
- J.-F. Chamayou and G. Letac (1999). Additive properties of the Dufresne laws and their multivariate extension. *J. Theoret. Probab.* **12**, 1045–1066.
- L. H. Y. Chen, L. Goldstein, and Q.-M. Shao (2011). *Normal approximation by Stein’s method*. Probability and its Applications (New York). Springer, Heidelberg. ISBN 978-3-642-15006-7.
- R. Durrett (2010). *Probability: theory and examples*. Cambridge University Press, 4th edition.

- E. Ford (2009). *Barabási-Albert random graphs, scale-free distributions and bounds for approximation through Steins method*. PhD thesis, University of Oxford.
- L. Goldstein (2007). L^1 bounds in normal approximation. *Annals of Probability* **35**, 1888–1930.
- L. Goldstein (2009). Personal communication and unpublished notes. Stein workshop, January 2009, Singapore.
- L. Goldstein and G. Reinert (1997). Stein’s method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7**, 935–952.
- S. Janson (2006). Limit theorems for triangular urn schemes. *Probab. Theory Related Fields* **134**, 417–452.
- S. Janson (2010). Moments of Gamma type and the Brownian supremum process area. *Probab. Surv.* **7**, 1–52.
- R. Lyons, R. Pemantle, and Y. Peres (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23**, 1125–1138.
- T. F. Móri (2005). The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.* **14**, 339–348.
- A. G. Pakes and R. Khattree (1992). Length-biasing, characterizations of laws and the moment problem. *Austral. J. Statist.* **34**, 307–322.
- E. Peköz and A. Röllin (2011). New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39**, 587–608.
- E. Peköz, A. Röllin, and N. Ross (2010). Geometric approximation and the equilibrium distribution coupling. *Preprint*. Available at <http://arxiv.org/abs/1005.2774>.
- G. Reinert (2005). Three general approaches to Stein’s method. In *An introduction to Stein’s method*, volume 4 of *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, pages 183–221. Singapore Univ. Press, Singapore.
- S. Ross and E. Peköz (2007). *A second course in probability*. www.ProbabilityBookstore.com, Boston, MA.
- C. Stein (1986). *Approximate computation of expectations*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7. Institute of Mathematical Statistics, Hayward, CA.